PCT

RAW SEQUENCE LISTING DATE: 07/08/2004
PATENT APPLICATION: **US/10/500,530** TIME: 16:00:43

Input Set : **A:\B45292SEQLIST.TXT**
Output Set: **N:\CRF4\07082004\J500530.raw**

ENTERED

(ps.6)

```
 4 <110> APPLICANT: Cindy Castado
 5       Joelle Thonnard
 7 <120> TITLE OF INVENTION: Novel Compounds
10 <130> FILE REFERENCE: B45292
C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/500,530
C--> 12 <141> CURRENT FILING DATE: 2004-07-01
12 <150> PRIOR APPLICATION NUMBER: GB 0200025.5
13 <151> PRIOR FILING DATE: 2002-01-02
15 <160> NUMBER OF SEQ ID NOS: 87
17 <170> SOFTWARE: FastSEQ for Windows Version 4.0
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 453
21 <212> TYPE: DNA
22 <213> ORGANISM: Haemophilus influenzae
24 <400> SEQUENCE: 1
25 gtgtgctatg agccatttat ttattaccca atgatgtgca atgaaaagat agcgcgtgct      60
26 attattcttg aagatgatgc gattgtatcg cacgaattcg aagcaattgt aaaagacagt     120
27 ttgaagaaag tttcaaaaaa tgttgaaatt ttatttttatg atcatggtaa agcaaaaagt    180
28 tattgctgga aaaaaacact tgtcaaaaat taccgtttag ttcactatcg taaaccctct    240
29 aaaacgtcta aacgtgcaat catgtgtaca acagcttatt taattacttt atctggcgct    300
30 caaaaactcc tacaaatagc ctatcctatc cgtatgcctg ctgactactt aactggtgct    360
31 ttacaattaa ctggactaaa ggcttatggt gttgaaccac cttgtgtatt taaaggcgca    420
32 atttcagaaa ttgatgcaat ggagcaacgc taa                                 453
34 <210> SEQ ID NO: 2
35 <211> LENGTH: 150
36 <212> TYPE: PRT
37 <213> ORGANISM: Haemophilus influenzae
39 <400> SEQUENCE: 2
40 Val Cys Tyr Glu Pro Phe Ile Tyr Tyr Pro Met Met Cys Asn Glu Lys
41  1               5                  10                  15
42 Ile Ala Arg Ala Ile Ile Leu Glu Asp Asp Ala Ile Val Ser His Glu
43             20                  25                  30
44 Phe Glu Ala Ile Val Lys Asp Ser Leu Lys Lys Val Ser Lys Asn Val
45         35                  40                  45
46 Glu Ile Leu Phe Tyr Asp His Gly Lys Ala Lys Ser Tyr Cys Trp Lys
47     50                  55                  60
48 Lys Thr Leu Val Lys Asn Tyr Arg Leu Val His Tyr Arg Lys Pro Ser
49 65                  70                  75                  80
50 Lys Thr Ser Lys Arg Ala Ile Met Cys Thr Thr Ala Tyr Leu Ile Thr
51                 85                  90                  95
52 Leu Ser Gly Ala Gln Lys Leu Leu Gln Ile Ala Tyr Pro Ile Arg Met
53             100                 105                 110
54 Pro Ala Asp Tyr Leu Thr Gly Ala Leu Gln Leu Thr Gly Leu Lys Ala
```

RAW SEQUENCE LISTING DATE: 07/08/2004
PATENT APPLICATION: US/10/500,530 TIME: 16:00:43

Input Set : A:\B45292SEQLIST.TXT
Output Set: N:\CRF4\07082004\J500530.raw

```
55          115                 120                 125
56 Tyr Gly Val Glu Pro Pro Cys Val Phe Lys Gly Ala Ile Ser Glu Ile
57      130                 135                 140
58 Asp Ala Met Glu Gln Arg
59 145                 150
61 <210> SEQ ID NO: 3
62 <211> LENGTH: 1032
63 <212> TYPE: DNA
64 <213> ORGANISM: Haemophilus influenzae
66 <400> SEQUENCE: 3
67 atgaaattaa aaaataaatt acaaatgtta aggttgggtc taggcaaata tttccttgat      60
68 aaaaaaaacg gattaaacag aataacaaat gttcctagaa gcatcctctt cctccgccaa     120
69 gacggaaaaa ttggggatta tgtggtgagc tcatttgtat tccgtgagat aaaaaaattt     180
70 aatccccaca ttaaaattgg tgtaatttgt accaaacaaa atgcttatct ttttaaacaa     240
71 aatccatata tcgatcaact ttactatgta aaaaagaaaa gtattttgga ttacatcaaa     300
72 tgtggtctag caattcaaaa agaacaatat gatttagtga ttgatccgac gattatgatt     360
73 cgtaatcgcg atcttttact tttacgctta atcaatgcca agcattatat tggctaccaa     420
74 aaagccaatt atggtttatt taatattaat ctggagggac aatttcactt ttcggaactc     480
75 tataaactcg ccttagaaaa agtgaatatt acggtacaag atataagcta tgacatccca     540
76 tttgataagc aaagtgcggt cgaaatttct gaatttttgc agaaaaacca actagaaaag     600
77 tatattgcta ttaattttta tggtgctgca agaatcaaaa aagtaaacaa tgacaacatc     660
78 aaaaaatatt tagattatct cacgcaagtc cgcggaggaa aaaagctggt gctattaagc     720
79 tatcctgaag taacagagaa attaacacaa ttgtcagccg attatccgca tatttttgtc     780
80 catccaacaa ccaagatctt tcataccatt gaattgattc gccactgtga tcaattaatc     840
81 tctacagaca cgtctactgt acatattgct tcaggtttta ataaaccaat tattggtatt     900
82 tataaagaag atcctattgc gtttacacat tggcaaccca gaagtcgggc agaaacgcac     960
83 atacttttct ataaagaaaa tattaatgag ctctcacctg aacaaattga ccctgcatgg    1020
84 cttgtcaaat ag                                                       1032
86 <210> SEQ ID NO: 4
87 <211> LENGTH: 343
88 <212> TYPE: PRT
89 <213> ORGANISM: Haemophilus influenzae
91 <400> SEQUENCE: 4
92 Met Lys Leu Lys Asn Lys Leu Gln Met Leu Arg Leu Gly Leu Gly Lys
93  1               5                  10                  15
94 Tyr Phe Leu Asp Lys Lys Asn Gly Leu Asn Arg Ile Thr Asn Val Pro
95             20                  25                  30
96 Arg Ser Ile Leu Phe Leu Arg Gln Asp Gly Lys Ile Gly Asp Tyr Val
97          35                  40                  45
98 Val Ser Ser Phe Val Phe Arg Glu Ile Lys Lys Phe Asn Pro His Ile
99      50                  55                  60
100 Lys Ile Gly Val Ile Cys Thr Lys Gln Asn Ala Tyr Leu Phe Lys Gln
101 65                  70                  75                  80
102 Asn Pro Tyr Ile Asp Gln Leu Tyr Tyr Val Lys Lys Lys Ser Ile Leu
103                 85                  90                  95
104 Asp Tyr Ile Lys Cys Gly Leu Ala Ile Gln Lys Glu Gln Tyr Asp Leu
105             100                 105                 110
106 Val Ile Asp Pro Thr Ile Met Ile Arg Asn Arg Asp Leu Leu Leu Leu
107         115                 120                 125
```

RAW SEQUENCE LISTING DATE: 07/08/2004
PATENT APPLICATION: US/10/500,530 TIME: 16:00:43

Input Set : A:\B45292SEQLIST.TXT
Output Set: N:\CRF4\07082004\J500530.raw

```
108 Arg Leu Ile Asn Ala Lys His Tyr Ile Gly Tyr Gln Lys Ala Asn Tyr
109     130                 135                 140
110 Gly Leu Phe Asn Ile Asn Leu Glu Gly Gln Phe His Phe Ser Glu Leu
111 145                 150                 155                 160
112 Tyr Lys Leu Ala Leu Glu Lys Val Asn Ile Thr Val Gln Asp Ile Ser
113                 165                 170                 175
114 Tyr Asp Ile Pro Phe Asp Lys Gln Ser Ala Val Glu Ile Ser Glu Phe
115             180                 185                 190
116 Leu Gln Lys Asn Gln Leu Glu Lys Tyr Ile Ala Ile Asn Phe Tyr Gly
117         195                 200                 205
118 Ala Ala Arg Ile Lys Lys Val Asn Asn Asp Asn Ile Lys Lys Tyr Leu
119     210                 215                 220
120 Asp Tyr Leu Thr Gln Val Arg Gly Gly Lys Lys Leu Val Leu Leu Ser
121 225                 230                 235                 240
122 Tyr Pro Glu Val Thr Glu Lys Leu Thr Gln Leu Ser Ala Asp Tyr Pro
123                 245                 250                 255
124 His Ile Phe Val His Pro Thr Thr Lys Ile Phe His Thr Ile Glu Leu
125             260                 265                 270
126 Ile Arg His Cys Asp Gln Leu Ile Ser Thr Asp Thr Ser Thr Val His
127         275                 280                 285
128 Ile Ala Ser Gly Phe Asn Lys Pro Ile Ile Gly Ile Tyr Lys Glu Asp
129     290                 295                 300
130 Pro Ile Ala Phe Thr His Trp Gln Pro Arg Ser Arg Ala Glu Thr His
131 305                 310                 315                 320
132 Ile Leu Phe Tyr Lys Glu Asn Ile Asn Glu Leu Ser Pro Glu Gln Ile
133                 325                 330                 335
134 Asp Pro Ala Trp Leu Val Lys
135             340
137 <210> SEQ ID NO: 5
138 <211> LENGTH: 813
139 <212> TYPE: DNA
140 <213> ORGANISM: Haemophilus influenzae
142 <400> SEQUENCE: 5
143 atgccagaat tacctgaagt tgaaaccaca aaaaatggaa ttagccctta tcttgaaggg      60
144 gctatcattg aaaaaattgt tgttcgccaa ccgaaattac gctggatggt aagcgaagaa     120
145 ttagcgcaaa ttacacaaca aaaagtcatc gcattaagtc gccgtgcgaa gtatttaatt     180
146 atccaacttg aaacaggcta tatgattgga catttaggga tgtcagggtc attgagagtt     240
147 gtggagaaag gggatcttat tgataaacat gatcatcttg atatcgtagt gaataacgga     300
148 aaagttgtgc gttataacga tcctcgtcgt tttggagcgt ggttatggac agagaagttg     360
149 aacgaatttc ctcttttttct gaaattaggc ccagagcctc tgtctgagga atttgattct     420
150 gattacttgt ggcaaaaaag tcgtaaaaaa cagaccgcac ttaaaacttt tttaatggat     480
151 aatgctgtcg tcgttggcgt tgggaatatc tatgcgaatg aaacgttatt tctttgtaac     540
152 ctacatccgc aaaaaacagc agggagttta actaaggcac aatgtgggca gttagtagaa     600
153 caaataaaac aagtgctgtc taacgcaatc caacaaggtg gtacgacgct aaaagatttt     660
154 ctccaaccgg atgggcgtcc aggctatttt gtccaagaat tgcgggttta tggtaataag     720
155 gataagcctt gtccaacatg tggcacaaaa atagaaagtt tagtgatagg gcaacgaaat     780
156 agtttctatt gccccaagtg tcagaagaga taa                                  813
158 <210> SEQ ID NO: 6
159 <211> LENGTH: 270
```

RAW SEQUENCE LISTING DATE: 07/08/2004
PATENT APPLICATION: US/10/500,530 TIME: 16:00:43

Input Set : A:\B45292SEQLIST.TXT
Output Set: N:\CRF4\07082004\J500530.raw

<212> TYPE: PRT
<213> ORGANISM: Haemophilus influenzae
<400> SEQUENCE: 6

Met	Pro	Glu	Leu	Pro	Glu	Val	Glu	Thr	Thr	Lys	Asn	Gly	Ile	Ser	Pro
1				5				10						15	
Tyr	Leu	Glu	Gly	Ala	Ile	Ile	Glu	Lys	Ile	Val	Val	Arg	Gln	Pro	Lys
		20					25						30		
Leu	Arg	Trp	Met	Val	Ser	Glu	Glu	Leu	Ala	Gln	Ile	Thr	Gln	Gln	Lys
	35					40					45				
Val	Ile	Ala	Leu	Ser	Arg	Arg	Ala	Lys	Tyr	Leu	Ile	Ile	Gln	Leu	Glu
50					55					60					
Thr	Gly	Tyr	Met	Ile	Gly	His	Leu	Gly	Met	Ser	Gly	Ser	Leu	Arg	Val
65				70					75					80	
Val	Glu	Lys	Gly	Asp	Leu	Ile	Asp	Lys	His	Asp	His	Leu	Asp	Ile	Val
			85					90					95		
Val	Asn	Asn	Gly	Lys	Val	Val	Arg	Tyr	Asn	Asp	Pro	Arg	Arg	Phe	Gly
		100					105					110			
Ala	Trp	Leu	Trp	Thr	Glu	Lys	Leu	Asn	Glu	Phe	Pro	Leu	Phe	Leu	Lys
	115					120					125				
Leu	Gly	Pro	Glu	Pro	Leu	Ser	Glu	Glu	Phe	Asp	Ser	Asp	Tyr	Leu	Trp
130					135					140					
Gln	Lys	Ser	Arg	Lys	Lys	Gln	Thr	Ala	Leu	Lys	Thr	Phe	Leu	Met	Asp
145				150					155					160	
Asn	Ala	Val	Val	Val	Gly	Val	Gly	Asn	Ile	Tyr	Ala	Asn	Glu	Thr	Leu
			165					170					175		
Phe	Leu	Cys	Asn	Leu	His	Pro	Gln	Lys	Thr	Ala	Gly	Ser	Leu	Thr	Lys
		180					185					190			
Ala	Gln	Cys	Gly	Gln	Leu	Val	Glu	Gln	Ile	Lys	Gln	Val	Leu	Ser	Asn
	195					200					205				
Ala	Ile	Gln	Gln	Gly	Gly	Thr	Thr	Leu	Lys	Asp	Phe	Leu	Gln	Pro	Asp
210					215					220					
Gly	Arg	Pro	Gly	Tyr	Phe	Val	Gln	Glu	Leu	Arg	Val	Tyr	Gly	Asn	Lys
225				230					235					240	
Asp	Lys	Pro	Cys	Pro	Thr	Cys	Gly	Thr	Lys	Ile	Glu	Ser	Leu	Val	Ile
			245					250					255		
Gly	Gln	Arg	Asn	Ser	Phe	Tyr	Cys	Pro	Lys	Cys	Gln	Lys	Arg		
		260					265					270			

<210> SEQ ID NO: 7
<211> LENGTH: 726
<212> TYPE: DNA
<213> ORGANISM: Haemophilus influenzae
<400> SEQUENCE: 7

atgagaattt tagccgcagg gagtttacgc cagcctttta cgttatggca acaagcatta 60
atccaacagt atcacctaca agtcgaaatt gaatttggac cggcggggtt gttgtgccaa 120
cgcattgagc aaggggaaaa agtggatttg tttgcctctg ccaatgatgc gcatcttagg 180
catttacaag cgcgatatcc tcatattcaa cttgtgcctt ttgctacaaa tcgtttatgt 240
ttaattgcaa agaaatcggt gattactcac catgatgaga attggttgac attattgatg 300
tcgccccact tacgcttagg agtatcgaca cctaaggcag atccttgtgg agattatact 360
ttggcattat tttcgaatat tgaaaaacgg catatgggct atggctcgga attaaaagaa 420

RAW SEQUENCE LISTING DATE: 07/08/2004
PATENT APPLICATION: US/10/500,530 TIME: 16:00:43

Input Set : A:\B45292SEQLIST.TXT
Output Set: N:\CRF4\07082004\J500530.raw

```
212 aaagcaatgg caatagttgg tggtccggat tctatcacta ttccaacagg acgaaatacc     480
213 gcagagtggc tttttgagca gaattatgct gatcttttca ttggttatgc gagtaatcat     540
214 caatctttgc gtcagcattc tgatatttgt gttttggata ttcctgatga gtataatgtg     600
215 agggcgaact atacattagc agcttttact gcggaagcat tacgccttgt ggactccttg     660
216 ctttgtttga cttgcggaca aaaatattta cgcgattgcg gctttttgcc tgccaatcat     720
217 agctga                                                               726
219 <210> SEQ ID NO: 8
220 <211> LENGTH: 241
221 <212> TYPE: PRT
222 <213> ORGANISM: Haemophilus influenzae
224 <400> SEQUENCE: 8
225 Met Arg Ile Leu Ala Ala Gly Ser Leu Arg Gln Pro Phe Thr Leu Trp
226  1               5                  10                  15
227 Gln Gln Ala Leu Ile Gln Gln Tyr His Leu Gln Val Glu Ile Glu Phe
228             20                  25                  30
229 Gly Pro Ala Gly Leu Leu Cys Gln Arg Ile Glu Gln Gly Glu Lys Val
230         35                  40                  45
231 Asp Leu Phe Ala Ser Ala Asn Asp Ala His Leu Arg His Leu Gln Ala
232     50                  55                  60
233 Arg Tyr Pro His Ile Gln Leu Val Pro Phe Ala Thr Asn Arg Leu Cys
234 65                  70                  75                  80
235 Leu Ile Ala Lys Lys Ser Val Ile Thr His His Asp Glu Asn Trp Leu
236                 85                  90                  95
237 Thr Leu Leu Met Ser Pro His Leu Arg Leu Gly Val Ser Thr Pro Lys
238             100                 105                 110
239 Ala Asp Pro Cys Gly Asp Tyr Thr Leu Ala Leu Phe Ser Asn Ile Glu
240         115                 120                 125
241 Lys Arg His Met Gly Tyr Gly Ser Glu Leu Lys Glu Lys Ala Met Ala
242     130                 135                 140
243 Ile Val Gly Gly Pro Asp Ser Ile Thr Ile Pro Thr Gly Arg Asn Thr
244 145                 150                 155                 160
245 Ala Glu Trp Leu Phe Glu Gln Asn Tyr Ala Asp Leu Phe Ile Gly Tyr
246                 165                 170                 175
247 Ala Ser Asn His Gln Ser Leu Arg Gln His Ser Asp Ile Cys Val Leu
248             180                 185                 190
249 Asp Ile Pro Asp Glu Tyr Asn Val Arg Ala Asn Tyr Thr Leu Ala Ala
250         195                 200                 205
251 Phe Thr Ala Glu Ala Leu Arg Leu Val Asp Ser Leu Leu Cys Leu Thr
252     210                 215                 220
253 Cys Gly Gln Lys Tyr Leu Arg Asp Cys Gly Phe Leu Pro Ala Asn His
254 225                 230                 235                 240
255 Ser
258 <210> SEQ ID NO: 9
259 <211> LENGTH: 741
260 <212> TYPE: DNA
261 <213> ORGANISM: Haemophilus influenzae
263 <400> SEQUENCE: 9
264 atgaatgaat tgagtttaga tgcagataag ctgttatttg gttatgataa gccgttgtat      60
265 ttaccactta ctttccaatg taagaaagga gaggttattt cggtatttgg aacaaatgga     120
```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 07/08/2004
PATENT APPLICATION: **US/10/500,530** TIME: 16:00:44

Input Set : **A:\B45292SEQLIST.TXT**
Output Set: **N:\CRF4\07082004\J500530.raw**

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:10; Xaa Pos. 163
Seq#:12; Xaa Pos. 80
Seq#:14; Xaa Pos. 49

VERIFICATION SUMMARY DATE: 07/08/2004
PATENT APPLICATION: **US/10/500,530** TIME: 16:00:44

Input Set : **A:\B45292SEQLIST.TXT**
Output Set: **N:\CRF4\07082004\J500530.raw**

L:12 M:270 C: Current Application Number differs, Replaced Current Application No
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:309 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:160
L:366 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:64
L:443 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:48